MZUMBE UNIVERSITY

FACULTY OF SCIENCE AND TECHNOLOGY

(FST)

DEPARTMENT OF COMPUTING SCIENCE STUDIES

(CSS)

PROGRAMME: BSc IN INFORMATION TECHNOLOGY AND SYSTEMS (ITS)
COURSE CODE: CSS 200
COURSE NAME: COLLABORATIVE NETWORK
INSTRUCTOR: DR. KISANJARA
TASK: GROUP ASSIGNMENT
SUBMISSION: MAY 13rd, 2022.

GROUP MEMBERS: GROUP NUMBER THREE (03)

S/N	STUDENT NAME	REGISTRATION NUMBER
1.	JOHARIA SEIF MAKULA	14322001/T.20
2.	GOODLUCK G. MWAKYUSA	14322007/T.20
3.	GEORGE E. SENG'ANA	14322008/T.20
4.	BERNADETHER K. SALVATORY	14322038/T.20

QUESTIONS

1. What is computational Grid? Why it needed? (Question 01)
2. Discuss applications of Grid computing. (Question 05)
3. Discuss Quality of Service in Grid Infrastructure. (Question 08)
4. Define Grid Computing. Explain Condor Middleware in detail(Question 12)
5. Discuss Authentication, authorization, and policy for grid computing (Question 14)
6. Discuss Grid computing models (Question 18)
7. Explain Planning and scheduling with reference to Condor (Question 20).
8. Explain real-time access to distributed instrumentation systems with example (Question 21)
9. Discuss Manageability, Monitorability and Integration for Grid (Question 22).
10. Discuss Scalability, Availability and Robustness for Grid (Question 24)

QUESTION NUMBER ONE (01)

Computational grid

Refers to a mechanism employed with processor architecture that combines diverse or various computer resources in distributed virtual organizations or from various domains to achieve its main objective.

In grid computing, the computers on the network can work on a task together, thus functioning as a supercomputer. It is designed to solve problems that are too big for a supercomputer while maintaining the flexibility to process numerous smaller problems. Computing grids deliver a multi-user infrastructure that accommodates the discontinuous demands of large information processing.

Computational grid is a computer network in which each **Computer's Resources** are shared with every other computer in the system. A grid computing system can be as simple as a collection of similar computers running on the same **Operating System**. The reason is to share resources, coordinated problem solving in dynamic and allowing multi-institutional virtual organization collaborations.

The following are the reasons why we need Computational grid:

i) High multi-processor speed.
Due to the collaboration of more than one processor it works in high speed in completing any task that performed at that time, so computational grid is very important in our life technology and collaborative organization.

ii) Proper exploitation of resources (underutilized resources).
This is the situation where by the resources of computers are inefficiently used in production process. A point lying below the production possibility curve underutilization of resources or inefficient utilization of resources.

iii) Parallel CPU capability.
Is a way in computing in which separates parts of an overall complex task, complex task are broken up and run simultaneously on multiple CPUs, thereby reducing the amount of time for processing.

iv) Sharing of Virtual resources.
Computational grid allocates to each virtual machine a share of physical resources. With the default resource, allocation settings all virtual machines associated with the same host receive an equal share of CPU per virtual CPU.

v) Enhance collaboration for virtual organizations.
It facilitate and increase collaboration among the organization even if are far away by getting work done when we are not co-located. It can mean working together to solve small problems, or collaborating to brainstorm innovative and creative solutions.

vi) Integration of different units.
A grid integrates and coordinates resources and users that live within different control domains grid technologies integrate different administrative units of the same company or different companies and address the issues of security, policy, payment, membership and so on.

vii) Ease data storage and management.

It makes effective management in ensuring organizations use storage resource effectively and which they store data securely in compliance with company policies and government regulations.

viii) Enhance business growth.

It helps to respond to market demand, increase market share and capitalize on growing brand. It often spurs innovation, helping to differentiate in market and stave off competition.

ix) For educational or intellectual purposes.

Computational grid is necessary and why it needed because it is used for learning in different institutes, colleges, schools and intellectual purposes.

QUESTION NUMBER FIVE (05)

Grid computing has various applications such as

i) Life Sciences

It encompasses all branches of information technologies such as in networking, databases, web application, mobile and desktop apps in performing different activities on team. Various life science disciplines such as computational biology, bioinformatics, genomics, neuroscience and other have embraced grid technology rapidly. Medical practitioners can access, collect and mine relevant data effectively. The grid also enables medical staff to perform large-scale simulations and analyses and connect remote instruments to existing medical infrastructures. For example, the MCell project explores cellular micro physiology using sophisticated diffusion and chemical reaction algorithms to simulate and study molecular interactions inside and outside cells. Grid technologies have enabled the large-scale deployment of various MCell modules, as MCell now runs on a large pool of resources, including clusters and supercomputers to perform biochemical simulations.

ii) Financial services

The need for computing power within the financial services industry has long been a critical issue. Increased competition, a changing regulatory environment and economic pressures have led many financial institutions to find that their existing computing resources are longer sufficient to meet demands. These demands include conducting complex tasks such as market analysis, pricing and risk management, which are essential if financial institutions are to outpace competitors and make the best possible decisions in the shortest space of time.

iii) Higher Education

Education infrastructure using grid-computing technology is used to enhance quality e-learning platform used for open universities. It allows the students, faculties, searching resources and distributed repository service among universities from anywhere, anytime can be accessed.

iv) Engineering services

Grid computing offers a suitable platform for design engineers to collaboratively work together and share knowledge and expertise in addition to the computational and data facility that can be combined to bear on complex designs.

v) Scientific research collaboration (e-Science)

Universities and institutions participating in advanced research collaboration programs have an enormous amount of data to analyze and process. Some examples of these projects include data analysis work for high-energy physics experiments, genome sequence analysis in COVID-19-like scenarios and the development of earth system models (ESM) by collecting data from several remote sensing sources. Organizations involved in research collaboration require substantial storage space as they regularly generate petabytes of data. They also need advanced computational resources to perform data-intensive processing. In this case, grid computing provides a resource-sharing mechanism by offering a single virtual organization that shares computing capabilities. The virtual supercomputer facilities the on-demand sharing of resources and integrates a secure framework for easy data access and interchange.

vi) Commercial applications

Grid computing supports various commercial applications such as the online gaming and entertainment industry, where computation-intensive resources such as computers and storage networks are essential. The resources are selected based on computing requirements in a gaming grid environment. It considers aspects such as the volume of traffic and the number of participating players.

OTHER APPLICATIONS

(i) Enabling partitioning process that involves breaking the problem into discrete pieces.
(ii) Discovering and scheduling of tasks and workflow.
(iii) Distributing problem data where necessary in data communications.
(iv) Providing and distributing application codes to specific system nodes.
(v) Results management assisting in the decision processes of the grid environment.
(vi) Implementation of autonomic features such as self-configuration, self-optimization, self-recovery, and self-management

Exploring some of these Grid applications and their usage patterns. We start with schedulers, which form the core component in most of the computational grids and follow up with other patterns.

Schedulers

Schedulers are types of applications responsible for the management of jobs, such as allocating
resources needed for any specific job, partitioning of jobs to schedule parallel execution of tasks,
data management, event correlation, and service-level management capabilities.

Resource-Broker

The resource broker provides *pairing* services between the service requester and the service provider. The pairing process in a resource broker involves allocation and support functions such as

• Allocating the appropriate resource or a combination of resources for the task execution

• Supporting users' deadline and budget constraints for scheduling optimizations

Load-Balancing

The Grid Computing infrastructure load-balancing issues are concerned with the traditional
load-balancing distribution of workload among the resources in a Grid Computing

environment.
This load-balancing feature must always be integrated into any system in order to avoid processing delays and over-commitment of resources. These kinds of applications can be built in connection with schedulers and resource managers. This level of load balancing involves partitioning of jobs, identifying the resources, and queuing the available jobs.

Grid-Portals

Grid portals provide uniform access to the grid resources. For example, grid portals provide capabilities for Grid Computing resource authentication, remote resource access, scheduling capabilities, and monitoring status information. Which help to reduce tasks complexity.

Integrated Solutions

These integrated solutions are a combination of the existing advanced middleware and application functionalities, combined to provide more coherent and high performance results across the Grid Computing environment. Integrated Grid Computing solutions will have more enhanced features to support more complex utilization of grids such as coordinated and optimized resource sharing, enhanced security management and cost optimizations, and areas yet to be explored.

QUESTION NUMBER EIGHT (08)

Quality of Service in Grid Infrastructure is the ability to provide the good and satisfactory grid services necessary for computational grid users (such as end-users) community. The Quality of Service (QoS) validations must be a basic feature in any Grid system and must be done concerning with the available resource matrices. These QoS features can be (for example) response time measures, aggregated performance, security fulfillment, resource scalability, and availability, autonomic features such as event correlation and configuration management, and partial fail over mechanisms.

QUESTION NUMBER TWELVE (12)

Condor is a resource management and job scheduling system. Condor middleware primary objective is to couple commodity machines in order to achieve the equivalent power of supercomputers with a significantly less expensive cost. Condor provides a job management mechanism, scheduling policy, priority scheme, resource monitoring, and resource management. The key feature of Condor is the ability to scavenge and manage wasted CPU power from idle desktop workstations across an entire virtual organization.

Grid middleware contain functionality for:

- ✓ Resource discovery, often through specialized information systems.
- ✓ Job submission, monitoring, and management.
- ✓ Authentication and authorization of users.

Additionally, middleware and related systems can incorporate solutions for advanced functionality such as resource brokering, accounting, and Grid-wide load balancing.

QUESTION NUMBER FOURTEEN (14)

Authentication-Deals with identifying various users of your grid or resources. Sets mechanism for the intended grid users to access the computational grid environment.

Or

Is the process of identifying securely the services and users that want to access the sites.

Authorization- Deals with determining whether a certain operation has met all protocols or rules set.

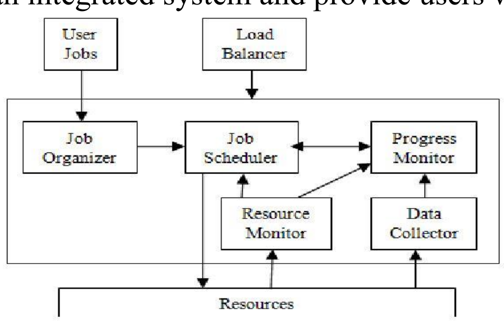
Authorization - allows for verifying the permissions to access specific resources (for instance data) when being authenticated on a site. Several access control mechanisms have been developed for Grids. Such as Encryption secures the content of the data. It is recognized that access control does not protect against all the security threats.

Policy - Allows overall alignment of organization rules and strategies in accessing or using the grid technologies. In distributed multi-institute environment like the GRID, each participating institute may want to enforce some limits on how its resources are being used by other institutes based on resource management in the case of single institute and how policy server based on GRID concepts based on developed, and then to show how to extend the policy server to support virtual organizations.

QUESTION NUMBER EIGHTEEN (18)

Grid computing model is a special kind computing model in which different computing devices, different computers within the same network share one or more resources. In the ideal grid computing system, every resource is shared, turning a computer network into a powerful supercomputer. They act as a base upon which grid computing relies to work nicely, such as:

i) Distributed load-balancing model.
Distributed load-balancing model, which can represent any grid topology into a forest structure. Load balancing strategy can be demonstrated at two levels: its principal objectives and the reduction of average response time of task and their transferring cost. The proposed strategy is naturally distributed with a local decision, which allows the possibility of avoiding use of wide area communication network. This model also integrates servers, storage systems and networks distributed within the network to form an integrated system and provide users with powerful computing and storage capacity.



ii) Economic model
Economic model deals with the costs of operating on a grid-computing environment. Such as hardware devices costs, software costs and other related costs like scheduling.

iii) Behavioral model
Behavioral model deals with how grid-computing environment behave in terms of handling various grid requirements on resources across the diverse community of the grid users.

QUESTION NUMBER TWENTY (20)

For Scheduling

Condor - new opted system for distributed computing. In contrast to dominant centralized control systems, condor is unique in its insistence that every participant in the system remain free to contribute as much as or as little as it cared to.

Condor performs low-penalty preemptive resume scheduling. Such as controlling system jobs.

When running jobs on remote machines, condor can often preserve the local execution environment via remote system calls. Where its mechanism for redirecting all of a jobs Input Output (I/O) related system calls back to the machines that submitted the job. Hence, allowing users to have no need to make data files available on remote workstation before condor execute their program there, even in the absence of a shared file system.

For planning

Condor can do more than effectively manage dedicated computer clusters. It can scavenge and manage wasted CPU power from otherwise the idle desktop workstation across an entire organization with minimal or low efforts.

For example, condor can be configured to run jobs on desktop workstation only when the keyboard and CPU are idle. If a job running on a workstation when the user returns and hit a certain key, condor can migrate the job to a different workstation and resume the job right where it left off.

Conclusively, condor uses matchmaking technology to bridge the gap between planning and scheduling. As it creates opportunities for planners and schedulers to work together while still respecting their essential independence.

QUESTION NUMBER 21

Real-time system

Refers to any information processing system, which has to respond to externally generated input stimuli within a known and specific period its correctness depends, on not only the logical result but also the time taken to deliver what was required.

There is a price to be paid for distributed systems. Network data traffic can cause congestion that actually slows performance. Data security and Integrity can also be more easily compromised in a distributed solution.

Examples

One of the best **examples** of a real-time system are those used to implement E-commerce such as **E-bay** and **Alibaba express**. If the network within ten milliseconds of being searched should deliver a chosen product, it can be considered to have met criteria of being a real-time process.

Some other examples of when using real-time systems would be beneficial are **Bank ATMs**, **air traffic control** and **fingerprint controlled devices**.

QUESTION NUMBER 22

Manageability

In grid computing manageability deals with overall arrangements and keeping track of all diverse resources shared across it such as data files, databases, operating system, scheduling, planning and other related matters.

Monitorability

In grid computing, monitor-ability deals with looking upon how the grid acts upon its users and how the users access the grid. Such as making sure of scalability within the grid by implementing various protocols and overlooking the grids performance such as accuracy of shared data files, response time and high processing speed.

Integration

In grid-computing integration is implemented to ensure proper data coordination between the existing advanced middleware and application functionalities, combined to provide more coherent and high performance results across the Grid Computing environment.

Integrated Grid Computing solutions will have more enhanced features to support more complex utilization of grids such as coordinated and optimized resource sharing, enhanced security management and cost optimizations, and areas yet to be explored.

QUESTION NUMBER 24

Scalability

This property is implemented on systems to handle a growing amount of work by adding resources. Grid application system requirements are likely to vary during application run-time, requiring underlying systems and infrastructure to access and scale computational, storage, and transfer capabilities on demand.

Availability

This property is implemented to ensure proper access of the grid to user whenever required. Grid applications and systems are typically composed through aggregation of computational resources, allowing Grids to exhibit very high levels of system availability despite system capacity varying over time due to resource volatility issues. Consistent levels of system access and quality of service improve the perception of Grid availability and stability.

Robustness

This property is implemented to ensure proper storage management and fast responsiveness of the grid systems. From a performance perspective, the construction of Grid systems is facilitated by improvements in computational and network capacity, and motivated by general availability of highly functional and well connected end systems. Increase in network capacity alone has led to changes in computing geometry and geography.

REFERENCES

1. D.P. Anderson (2004). BOINC: A system for public-resource computing and storage. In 5th IEEE/ACM International Workshop on Grid Computing, pages 4–10.

2. Foster, I and C. Kesselman (1999). The Grid: Blueprint for new computing infrastructure, San Francisco: Morgan Kaufmann publisher.

3. S. Venugopal, R. Buyya and L. Winton (2004). A grid service broker for scheduling distributed data-oriented application on global Grids.

4. International Technical Support Organization SG24-6778-00 Introduction to Grid Computing December 2005.

5. https://fardapaper.ir/mohavaha/uploads/2017/11/Grid-Computing-and-Security-Issues.pdf

6. https://docs.huihoo.com/globus/gt3-tutorial/ch01s03.html

7. https://en.wikipedia.org/wiki/Semantic_grid

8. https://www.quora.com/What-are-the-grid-computing-models